

## Bibliographic Fields

## Document Identity

(19)【発行国】  
日本国特許庁 (JP)  
(12)【公報種別】  
公開特許公報 (A)  
(11)【公開番号】  
特開 2000-132559 (P2000-132559 A)  
(43)【公開日】  
平成 12 年 5 月 12 日 (2000. 5. 12)

(19) [Publication Office]  
Japan Patent Office (JP)  
(12) [Kind of Document]  
Unexamined Patent Publication (A)  
(11) [Publication Number of Unexamined Application]  
Japan Unexamined Patent Publication 2000-132559 (P2000-132559A)  
(43) [Publication Date of Unexamined Application]  
2000 May 12\* (2000.5.12)

## Public Availability

(43)【公開日】  
平成 12 年 5 月 12 日 (2000. 5. 12)

(43) [Publication Date of Unexamined Application]  
2000 May 12\* (2000.5.12)

## Technical

(54)【発明の名称】  
情報フィルタリングシステムにおけるプロファイル更新方法及び情報フィルタリングシステム  
(51)【国際特許分類第 7 版】  
G06F 17/30  
【FI】  
G06F 15/403 340 A  
15/40 310 F  
15/403 350 C  
【請求項の数】  
22  
【出願形態】  
OL  
【全頁数】  
14  
【テーマコード(参考)】  
5B075  
【F ターム(参考)】  
5B075 KK07 ND03 NK02 NK35 NR02 NR1  
2 PR06 PR08 QM08

(54) [Title of Invention]  
profile update method and data filtering system in data filtering system  
(51) [International Patent Classification, 7th Edition]  
G06F17/30  
[FI]  
G06F15/403340A  
15/40310F  
15/403350C  
[Number of Claims]  
22  
[Form of Application]  
OL  
[Number of Pages in Document]  
14  
[Theme Code (For Reference)]  
5B075  
[F Term (For Reference)]  
5B075KK07ND03NK02NK35NR02NR12PR06PR08QM08

**Filing**

【審査請求】

[Request for Examination]

未請求

Unrequested

(21)【出願番号】

(21) [Application Number]

特願平10-301992

Japan Patent Application Hei 10- 301992

(22)【出願日】

(22) [Application Date]

平成10年10月23日(1998. 10. 23)

1998 October 23\* (1998.10.23)

**Parties****Applicants**

(71)【出願人】

(71) [Applicant]

【識別番号】

[Identification Number]

000005108

000005108

【氏名又は名称】

[Name]

株式会社日立製作所

Hitachi Ltd. (DB 69-054-1503)

【住所又は居所】

[Address]

東京都千代田区神田駿河台四丁目6番地

Tokyo Chiyoda-ku Kanda Surugadai 4-Chome 6

**Inventors**

(72)【発明者】

(72) [Inventor]

【氏名】

[Name]

森本 康嗣

Morimoto Yasushi

【住所又は居所】

[Address]

東京都国分寺市東恋ヶ窪一丁目280番地 株式会社日立製作所中央研究所内

Tokyo Kokubunji City Higashi Koigakubo 1-280 Hitachi Ltd. (DB 69-054-1503) Central Research Laboratory \*

(72)【発明者】

(72) [Inventor]

【氏名】

[Name]

三科 雄介

Mishina Yusuke

【住所又は居所】

[Address]

東京都国分寺市東恋ヶ窪一丁目280番地 株式会社日立製作所中央研究所内

Tokyo Kokubunji City Higashi Koigakubo 1-280 Hitachi Ltd. (DB 69-054-1503) Central Research Laboratory \*

(72)【発明者】

(72) [Inventor]

【氏名】

[Name]

梶 博行

\*Hiroyuki

【住所又は居所】

[Address]

東京都国分寺市東恋ヶ窪一丁目280番地 株式会社日立製作所中央研究所内

Tokyo Kokubunji City Higashi Koigakubo 1-280 Hitachi Ltd. (DB 69-054-1503) Central Research Laboratory \*

## Agents

(74)【代理人】

【識別番号】

100068504

【弁理士】

【氏名又は名称】

小川 勝男

## Abstract

(57)【要約】

【課題】

情報フィルタリング用のプロフィールを精密化する。

【解決手段】

ユーザの興味を設定したプロフィールを用いて入力情報を適切なユーザに配信する情報フィルタリングシステムにおいて、類似した興味を持つユーザのプロフィールを用いることにより自分のプロフィールを精密化する。

さらに、類似した興味を持つ他のユーザを自動的に決定することにより、効率的なプロフィール精密化を行う。



## Claims

【特許請求の範囲】

【請求項 1】

プロフィールに記述された情報要求と入力された電子化文書とを比較し、上記情報要求を満たす電子化文書を上記プロフィールのユーザに配信する情報フィルタリングシステムにおけるプロフィールの更新方法において、

上記プロフィールごとに上記配信された電子化文書の情報を参照履歴情報として保存し、

第一のプロフィールについて、その参照履歴情報が上記保存された他のプロフィールのうちその参照履歴情報との類似度が所定値よりも高い他のプロフィールを特定し、

上記特定された他のプロフィールに基づき上記

(74) [Attorney(s) Representing All Applicants]

[Identification Number]

100068504

[Patent Attorney]

[Name]

Ogawa Katsuo

(57) [Abstract ]

[Problems to be Solved by the Invention ]

profile for data filtering to precision is converted.

[Means to Solve the Problems ]

To precision you convert your own profile by using profile of the user which has interest which resembles in data filtering system which transmits input information to appropriate user making use of profile which sets interest of user .

Furthermore, it converts efficient profile precision by deciding other user which has interest which resembles in automatic .

[Claim (s )]

[Claim 1 ]

In update method of profile in data filtering system which transmits electronic document which information need which is described to profile compares electronic document which is inputted, fills up above-mentioned information need to user of the above-mentioned profile ,

Description above is transmitted it retains data of electronic document which in every above-mentioned profile as reference history data ,

Concerning profile of first , reference history data description above similarity of reference history data among other profile which are retained specific does high other profile in comparison with the specified value , profile update method . which designates that profile of

第一のプロファイルを更新することを特徴とする  
プロフィール更新方法。

【請求項 2】

請求項 1 記載のプロファイル更新方法において、

上記参照履歴情報として、上記プロファイルに基づき配信された電子化文書の文書識別子と上記配信された電子化文書に対するユーザの興味との適合を示す適合度とを含み、

上記第一のプロファイルの参照履歴情報より得られる電子化文書のうち高い適合度が付与された電子化文書の集合と上記他のプロファイルの参照履歴より得られる電子化文書のうち高い適合度が付与された電子化文書の集合との重なりを上記類似度とすることを特徴とするプロファイル更新方法。

【請求項 3】

請求項 2 記載のプロファイル更新方法において、

上記適合度はユーザによる入力され、または配信された電子化文書のうちユーザにより全文が閲覧されたものを上記適合度が高いと擬制することを特徴とするプロファイル更新方法。

【請求項 4】

請求項 2 記載のプロファイル更新方法において、

上記第一のプロファイルのユーザに配信された電子化文書(A)に対する適合度を、上記第一のプロファイルのユーザと類似した興味を有するユーザが上記電子化文書(A)に付した適合度により擬制することを特徴とするプロファイル更新方法。

【請求項 5】

請求項 1 記載のプロファイル更新方法において、

上記他のプロファイルとして、所定のプロファイルの一つのプロファイルとしてまとめたものを含むことを特徴とするプロファイル更新方法。

【請求項 6】

プロファイルに記述された情報要求と入力され

above-mentioned first is renewed on basis of above-mentioned other profile which specific is done as feature

[Claim 2]

In profile update method which is stated in Claim 1,

Including fidelity conformity of document identifier of electronic document which is transmitted as above-mentioned reference history data, on basis of the above-mentioned profile and description above interest of user for electronic document which is transmitted is shown,

From reference history data of profile of above-mentioned first from reference history of gathering and above-mentioned other profile of electronic document where fidelity among electronic document which are acquired it is high is granted profile update method, which designates that overlap of gathering of electronic document where fidelity among electronic document which are acquired it is high is granted is designated as above-mentioned similarity as feature

[Claim 3]

In profile update method which is stated in Claim 2,

Above-mentioned fidelity is inputted with user, profile update method, which designates that those where full text is perused or by inside user of electronic document which is transmitted are controlled when above-mentioned fidelity is high, pseudo as feature

[Claim 4]

In profile update method which is stated in Claim 2,

Fidelity for electronic document (A) which is transmitted in user of the profile of above-mentioned first, with fidelity user which possesses interest which resembles user of profile of the above-mentioned first attaches on above-mentioned electronic document (A) pseudo profile update method, which designates that it controls as feature

[Claim 5]

In profile update method which is stated in Claim 1,

As above-mentioned other profile, to with predetermined profile as profile of one profile update method, which designates that those which were collected are included as feature

[Claim 6]

In update method of profile in data filtering system which

た電子化文書とを比較し、上記情報要求を満たす電子化文書を上記プロフィールのユーザに配信する情報フィルタリングシステムにおけるプロフィールの更新方法において、

第一のプロファイルの情報要求と他のプロフィールの情報要求との差分である差分タームを抽出し、

上記差分タームを上記第一のプロファイルにマージすることを特徴とするプロフィール更新方法。

#### 【請求項 7】

請求項 6 記載のプロファイル更新方法において、

上記他のプロフィールは、上記第一のプロファイルのユーザにより指定される、もしくは上記第一のプロファイルの情報要求が上記情報フィルタリングシステムに登録されている他のプロフィールの情報要求との類似度が所定値よりも高いことにより特定されることを特徴とするプロフィール更新方法。

#### 【請求項 8】

請求項 6 記載のプロファイル更新方法において、

上記抽出された差分タームは、上記第一のプロファイルにマージする差分タームの指定をユーザにより受けるため、表示画面上に表示されることを特徴とするプロフィール更新方法。

#### 【請求項 9】

請求項 8 記載のプロファイル更新方法において、

上記他のプロフィールのユーザに配信されており、かつ上記第一のプロファイルのユーザには配信されていない電子化文書であって、上記差分タームを含むものを支援情報として抽出し、

上記抽出された支援情報は、ユーザによる差分タームの指定を支援するため、上記表示画面上に表示されることを特徴とするプロフィール更新方法。

#### 【請求項 10】

請求項 9 記載のプロファイル更新方法において、

上記情報フィルタリングシステムは、各プロファ

transmits electronic document which information need which is described to profile compares electronic document which isinputted, fills up above-mentioned information need to user of theabove-mentioned profile ,

difference [taamu ] which is a difference of information need of profile of first and information need of other profile is extracted,

Above-mentioned difference [taamu ] in profile of above-mentioned first merge profile update method , which designates that it does as feature

[Claim 7 ]

In profile update method which is stated in Claim 6,

Above-mentioned other profile is appointed by user of profile of above-mentioned first , or similarity of information need of other profile where information need of profile of above-mentioned first isregistered to above-mentioned data filtering system in comparison with specified value due to high especially specific profile update method . which designates that it isdone as feature

[Claim 8 ]

In profile update method which is stated in Claim 6,

Description above as for difference [taamu ] which is extracted, in order in profile of above-mentioned first with user appointmentof difference [taamu ] which merge is done to receive, profile update method . whichdesignates that it is indicated on display screen as feature

[Claim 9 ]

In profile update method which is stated in Claim 8,

We are transmittd in user of above-mentioned other profile , atsame time with electronic document which is not transmittd, we extract thosewhich include above-mentioned difference [taamu ] in user of profile of above-mentioned first as supporting information ,

Description above as for supporting information which is extracted, in order tosupport appointment of difference [taamu ] with user , profile update method . which designates that it is indicated on above-mentioned display screen asfeature

[Claim 10 ]

In profile update method which is stated in Claim 9,

Above-mentioned data filtering system keeps user data

イルのユーザに関するユーザ情報をあらかじめ保持し、

上記他のプロフィールのユーザのユーザ情報を支援情報として抽出し、

上記抽出された支援情報は、ユーザによる差分タームの指定を支援するため、上記表示画面上に表示されることを特徴とするプロフィール更新方法。

【請求項 11】

請求項 10 記載のプロフィール更新方法において、

上記支援情報として抽出されるユーザ情報は、一部制限可能であることを特徴とするプロフィール更新方法。

【請求項 12】

プロフィールに記述された情報要求と入力された電子化文書とを比較し、上記情報要求を満たす電子化文書を上記プロフィールのユーザに配信する情報フィルタリングシステムにおけるプロフィールの更新方法において、

第一のプロフィールの更新起動タイミングを決定し、

上記更新起動タイミングで、第一のプロフィールの情報要求と所定のプロフィールの情報要求との差分である差分タームを抽出し、

上記差分タームにより、上記第一のプロフィールを更新することを特徴とするプロフィール更新方法。

【請求項 13】

請求項 12 記載のプロフィール更新方法において、

上記所定のプロフィールは、上記第一のユーザによって指定された他のユーザのプロフィール、または上記第一のプロフィールの情報要求が類似する他のユーザのプロフィールであることを特徴とするプロフィール更新方法。

【請求項 14】

請求項 12 記載のプロフィール更新方法において、

上記プロフィールごとに上記プロフィールに基づき配信された電子化文書の文書識別子と上記配信された電子化文書に対するユーザの興味との適合を示す適合度を含む参照履歴情報

regarding user of each profile beforehand,

It extracts user data of user of above-mentioned other profile as supporting information ,

Description above as for supporting information which is extracted, in order to support appointment of difference [taamu ] with user , profile update method , which designates that it is indicated on above-mentioned display screen as a feature

[Claim 11 ]

In profile update method which is stated in Claim 10,

As for user data which is extracted as above-mentioned supporting information , part restriction profile update method , which designates that it is possible as a feature

[Claim 12 ]

In update method of profile in data filtering system which transmits electronic document which information need which is described to profile compares electronic document which is inputted, fills up above-mentioned information need to user of the above-mentioned profile ,

Renewal starting timing of profile of first is decided,

With above-mentioned renewal starting timing , difference [taamu ] which is a difference of information need of profile of first and information need of the predetermined profile is extracted,

With above-mentioned difference [taamu ], profile update method , which designates that the profile of above-mentioned first is renewed as a feature

[Claim 13 ]

In profile update method which is stated in Claim 12,

As for above-mentioned predetermined profile , profile update method , which designates that it is a profile of other user to which profile , of other user which is appointed with user of above-mentioned first or the information need of profile of above-mentioned first resembles as a feature

[Claim 14 ]

In profile update method which is stated in Claim 12,

Reference history data which includes fidelity conformity of document identifier of electronic document which is transmitted on basis of above-mentioned profile and description above interest of user for electronic document

を保存し、

上記所定のプロフィールは、上記第一のプロファイルとその参照履歴情報が類似するプロファイルであることを特徴とするプロフィール更新方法。

#### 【請求項 15】

請求項 12 記載のプロファイル更新方法において、

上記各プロフィールの更新状況を監視し、上記所定のプロフィールの更新処理がなされたことが検出された時点で更新起動を決定することを特徴とするプロフィール更新方法。

#### 【請求項 16】

プロフィールに記述された情報要求と電子化文書とを比較し、上記プロフィールの情報要求を満たす電子化文書を判定する情報フィルタ機能と、

第一のプロファイルと第二のプロファイルとの類似度を判定するプロフィール類似判定機能と、

上記プロフィール類似判定機能により上記第二のプロファイルが上記第一のプロファイルに類似すると判定された場合、上記第二のプロファイルの情報要求により上記第一のプロファイルの情報要求を更新するプロフィール更新機能とを有することを特徴とするコンピュータ読取可能な記録媒体。

#### 【請求項 17】

請求項 17 記載のコンピュータ読取可能な記録媒体において、

上記情報フィルタ機能により上記プロフィールの情報要求を満たすと判定された電子化文書に対して、ユーザの興味との適合の度を合示す適合度を入力する適合度入力機能を有し、

上記プロフィール類似判定機能において、上記適合度入力機能により入力された上記適合度により、上記第一のプロファイルと第二のプロファイルとの類似度を判定することを特徴とするコンピュータ読取可能な記録媒体。

#### 【請求項 18】

情報ネットワークに接続され、上記情報ネットワークを介して配送された電子化文書をフィルタリングする情報フィルタリングシステムにおいて、

which is transmitted is shown is retained in every above-mentioned profile ,

As for above-mentioned predetermined profile , profile update method . which designates that it is a profile of above-mentioned first and a profile to which reference history data resembles as feature

#### [Claim 15]

In profile update method which is stated in Claim 12,

profile update method . which designates that renewal starting is decided with the time point where watching renewal status of above-mentioned each profile , being able to do renewal treatment of above-mentioned predetermined profile is detected as feature

#### [Claim 16]

data filter function which decides electronic document which compares information need and electronic document which are described to profile , fills up information need of above-mentioned profile and,

profile of first and profile similar decision function which decides similarity of second profile and,

When it is decided, that above-mentioned second profile resembles to the profile of above-mentioned first , with above-mentioned profile similar decision function computer readable recording medium . which designates that it possesses the profile renewal function which renews information need of profile of the above-mentioned first with information need of above-mentioned second profile as feature

#### [Claim 17]

In computer readable recording medium which is stated in Claim 17,

Fidelity input function which inputs fidelity extent of conformity of interest of user is shown vis-a-vis electronic document which is decided that information need of above-mentioned profile is filled up with above-mentioned data filter function, possessing,

computer readable recording medium , which designates profile of above-mentioned first and that similarity of second profile is decided as feature with above-mentioned fidelity it is inputted in above-mentioned profile similar decision function, by above-mentioned fidelity input function

#### [Claim 18]

In data filtering system which electronic document which is connected by data network , through above-mentioned data network , is delivered filtering is done,

複数のプロフィールを格納するプロフィール記録手段と、

上記プロフィールの情報要求を満たす上記電子化文書を判定する情報フィルタプログラムと、上記複数のプロフィール間の類似度を判定するプロフィール類似判定プログラムと、プロフィールの情報要求をその類似するプロフィールの情報要求により更新するプロフィール更新プログラムとを格納するメモリ手段と、

上記メモリ手段に格納されたプログラムを実行する処理手段と、

上記類似判定プログラムを上記処理手段により実行して得られた各プロフィールに類似するプロフィール情報を格納する類似プロフィール情報記録手段とを有することを特徴とする情報フィルタリングシステム。

#### 【請求項 19】

請求項 18 の情報フィルタリングシステムにおいて、

上記メモリ手段は、上記プロフィールに基づき記憶された電子化文書に対するユーザの興味との適合を示す適合度の入力を受ける適合度入力プログラムを格納しており、

上記プロフィールごとに上記プロフィールに基づき記憶された電子化文書の文書識別子と上記適合度入力プログラムを上記処理手段により実行して得られた上記適合度とを含む参照履歴情報を格納する参照履歴記録手段とを有し、

上記参照履歴記録手段を参照し、第一のプロファイルにおいて高い適合度が付与された電子化文書の集合と第二のプロファイルにおいて高い適合度が付与された電子化文書の集合との重なりにより上記複数のプロフィール間の類似度を判定することを特徴とする情報フィルタリングシステム。

#### 【請求項 20】

情報プロバイダより情報ネットワークを介して電子化文書の配信を受け、上記電子化文書を各ユーザのプロファイルによりフィルタリングして、上記プロフィールに記述された情報要求を満たす電子化文書を配信する情報フィルタリング方法において、

所定のタイミングで第一のユーザに対して、そのプロフィールの更新起動を要請し、

profile recording means which houses profile of plural and,

memory means which houses profile similar decision program which decides similarity between profile of data filter program and above-mentioned plural which decide above-mentioned electronic document which fills up information need of theabove-mentioned profile and profile renewal program which renews the information need of profile that with information need of profile which resemblesand,

processing means which executes program which is housed in theabove-mentioned memory means and,

Executing above-mentioned similar decision program with theabove-mentioned processing means , data filtering system . which designates that it possessessimilar profile information recording means which houses profile data which resembles to each profile which it acquires as feature

#### 【Claim 19】

In data filtering system of Claim 18 ,

Above-mentioned memory means has housed fidelity input program whichreceives input fidelity conformity of interest of user for electronic document which is transmitted on basis of above-mentioned profile is shown,

Executing document identifier and above-mentioned fidelity input program of electronic document which is transmitted on basis of above-mentioned profile in every above-mentioned profile with above-mentioned processing means ,reference history recording means which houses reference history data which includesabove-mentioned fidelity it acquires possessing,

Above-mentioned reference history recording means is referred to, data filtering system . whichdesignates that similarity between profile of above-mentioned plural is decided gathering of electronic document where fidelity it is high with overlap of gathering of electronic document where fidelity it is high in the profile of first is granted and in second profile is granted as feature

#### 【Claim 20】

Through data network from data provider , to receive transmission of the electronic document , above-mentioned electronic document filtering doing with profile ofeach user , regarding to data filtering method which transmits electronic document whichfills up information need which is described to above-mentioned profile ,

Renewal starting of profile is requested with predetermined timing vis-a-vis user of first ,



上記要請に応じて、上記第一のユーザからのプロフィールの更新起動命令を受けた場合には、

上記プロフィールを第二のユーザのプロファイルに基づき更新することと特徴とする情報フィルタリング方法。

#### 【請求項 21】

請求項 20 記載の情報フィルタリング方法において、

上記第二のユーザは、上記第一のユーザによって指定された他のユーザ、または上記プロフィールの情報要求が類似するプロフィールを所有する他のユーザであることを特徴とする情報フィルタリング方法。

#### 【請求項 22】

請求項 20 記載の情報フィルタリング方法において、

上記プロフィールごとに上記プロフィールに基づき記信された電子化文書の文書識別子と上記記信された電子化文書に対するユーザの興味との適合を示す適合度とを含む参照履歴情報を保存し、

上記第二のユーザは、上記第一のユーザのプロファイルとその参照履歴情報が類似するプロフィールを有する他のユーザであることを特徴とする情報フィルタリング方法。

### Specification

#### 【発明の詳細な説明】

【0001】

#### 【発明の属する技術分野】

本発明は、情報フィルタリングシステムに関する。

特に、情報フィルタリングに用いられるユーザのプロファイルの精密化方法に関する。

【0002】

#### 【従来の技術】

現在、ネットワークの発達、特にインターネットの普及に伴って、ネットワークを介して入手可能な電子化文書の量が膨大になってきている。

インターネット上の電子化文書に対するアクセス方法は、従来、http(hyper text transfer protocol)による WWW(World Wide Web)文書のプ

According to above-mentioned request, when renewal starting command of profile from user of above-mentioned first is received,

data filtering method, which designates that above-mentioned profile is renewed on basis of profile of second user as feature

#### 【Claim 21】

Regarding to data filtering method which is stated in Claim 20,

As for above-mentioned second user, data filtering method, which designates that it is another user which owns profile to which other user, or the information need of above-mentioned profile which is appointed with user of above-mentioned first resembles as feature

#### 【Claim 22】

Regarding to data filtering method which is stated in Claim 20,

Reference history data which includes fidelity conformity of document identifier of electronic document which is transmitted on basis of above-mentioned profile and description above interest of user for electronic document which is transmitted is shown is retained in every above-mentioned profile,

As for above-mentioned second user, data filtering method, which designates that it is another user which possesses profile of user of the above-mentioned first and profile to which reference history data resembles as feature

#### 【Description of the Invention】

【0001】

#### 【Technological Field of Invention】

this invention regards data filtering system.

Especially, it regards precision conversion method of profile of the user which is used for data filtering.

【0002】

#### 【Prior Art】

Presently, attendant upon advancement of network and spread of especially Internet, through network, quantity of available electronic document has become expansion.

As for access method for electronic document on Internet, until recently, browsing of WWW (World Wide Web) document was center with http (hypertext transfer protocol).

ウジングが中心であった。

WWW 文書のブラウジングは、Web ブラウザと呼ばれるソフトウェア上で、その WWW 文書のアドレスである URL(Uniform Resource Locator)を指定することによって行う。

URL は、その WWW 文書が存在する計算機及びその計算機上でのディレクトリなどを一意に指定する。

ユーザが個別に WWW 文書を指定してアクセスするブラウジングは、通常、プル型と呼ばれている。

【0003】

これに対して、プッシュ型と呼ばれる情報配信型のサービス(ユーザは個別に WWW 文書を指定することなくアクセスする)が注目されている。

このような情報配信型のサービスでは、各ユーザが興味を持つ情報(WWW 文書)を配信することが重要である。

【0004】

情報フィルタリングシステムは、各ユーザが興味をもつ事柄についての情報を記述したデータ(インタレストプロファイル以下、「プロファイル」と呼ばれる)に基づいて、配信される情報をフィルタリングするシステムである。

情報フィルタリングシステムは、WWW 文書がネットワークを通じて配信されると、配信された WWW 文書と各ユーザのプロファイルとを比較することにより、その WWW 文書に興味を持つであろうユーザを決定し、そのユーザに対してのみにその WWW 文書を配送する。

これにより、ユーザは、自分が興味を持つ WWW 文書のみを受け取ることができる。

【0005】

したがって、情報フィルタリングにおいて最も大きな課題は、配信された WWW 文書がユーザの興味に合致するかどうかを正しく判定することであり、そのためには、プロファイルにユーザの興味を正確に記述されている必要がある。

プロファイルを精密化するために、特開平 9-153 064 号公報に開示されているレリバンスフィードバック技術が提案されている。

レリバンスフィードバックでは、プロファイルに基づいて各ユーザに配信された情報に対して、ユーザが適合度(ユーザの興味に対して配信され

On software which is called Web browser, it does browsing of the WWWdocument, by fact that URL (UniformResourceLocator) which is a address of WWWdocument isappointed.

URL appoints computer and directory etc on its computer where the WWWdocument exists to unique.

user appointing WWWdocument individually, browsing which access is done, usually, is called pull type.

【0003】

Vis-a-vis this, service (user without appointing WWWdocument individually access does) of data transmission type which is called push type is observed.

With service of data transmission type a this way, it is important to transmit data (WWWdocument) where each user has interest.

【0004】

data filtering system is system which data which is transmitted on the basis of data ([intarecutopurofairu] (Below, "profile") With it is called) which describes data concerning item where each user has interest, filtering is done.

user data filtering system, when WWWdocument it is transmitted via network, probably will have interest in WWWdocument WWWdocument which is transmitted and by comparing profile of each user, is decided, too only vis-a-vis user WWWdocument is delivered.

Because of this, to receive only WWWdocument where by his has interest it is possible user.

【0005】

Therefore, in data filtering, largest problem by fact that it decides correctly whether or not WWWdocument which is transmitted coincides to interest of user, for that, has necessity for interest of user to be described to profile accurately.

In order to precision to convert profile, [reribansufidobakku] technology which is disclosed in Japan Unexamined Patent Publication Hei 9-153064 disclosure has been proposed.

With [reribansufidobakku], user inputs fidelity (Extent data which is transmitted vis-a-vis interest of the user coinciding) vis-a-vis data which is transmitted in each user on basis of

た情報が合致する程度)を入力する。

システムは、入力された適合度をプロフィールにフィードバック(例えばプロフィールに付された重みを変更する)することによってプロフィールがユーザの興味をよりよく反映するように精密化する。

[0006]

また、別のアプローチとして特開平 9-265478 号公報に開示されている協調フィルタリング技術が提案されている。

このアプローチは自分と同じ興味を有する他のユーザを発見し、他のユーザに配信された情報は自分にも同様に配信してもらうことにより、興味のある WWW 文書をめを少なくして受け取るようにするものである。

[0007]

[発明が解決しようとする課題]

レリバンスフィードバック技術では、プロフィールの精密化をユーザ毎に行う。

しかし、各ユーザを個別に扱っている限り、精密化の精度には限界があった。

従来のレリバンスフィードバックでは、各ユーザが参照した文書にもとづいて自己のプロファイルの精密化を行う。

しかし、ユーザ各人が参照できる文書数には限りがあるため、プロフィール精密化の精度には限界がある。

[0008]

また、協調フィルタリング技術では、自己のプロファイル自体は修正されず、また、他人の興味は自分の興味と完全に一致していることはまれであるから配信される WWW 文書量の増大が避けられない。

[0009]

本発明の目的は、以上の従来技術の問題を踏まえて、新しいプロフィール精密化方法およびそれを適用した情報フィルタリングシステムを提供する。

本発明では、同じ情報フィルタリングシステムを利用する他のユーザの知識を利用して自己のプロファイルを精密化する。

[0010]

[課題を解決するための手段]

profile .

system , in order in profile feedback for profile to reflect interest of user well by fact that (weight which is attached on for example profile is modified ) , to precision converts fidelity it is inputted.

[0006]

In addition, cooperation filtering technology which is disclosed in Japan Unexamined Patent Publication Hei 9-265478disclosure asanother approach is proposed.

data where this approach as by your discover other user whichpossesses same interest , is transmitted in other user issomething which decreasing a leak, it tries to receive WWWdocument whichhas interest by you have transmitting in same way even by your.

[0007]

[Problems to be Solved by the Invention ]

With [reribansufiidobakku ] technology , it converts profile precision in every user .

But, if each user is handled individually, there was a limit in the precision of precision conversion.

With conventional [reribansufiidobakku ] , it converts profile of self precision on thebasis of document which each user refers to.

But, because there is limit in number of documents which user each personcan refer to, there is a limit in precision of profile precision conversion.

[0008]

In addition, with cooperation filtering technology , as for profile itself of self not to be corrected, in addition, because interest of othersagreeing with your own interest completely be fixed is you cannot avoidincrease of WWWdocument quantity which is transmitted.

[0009]

objective of this invention offers new profile precision conversion method and the data filtering system which applies that on basis of problem of Prior Art above.

With this invention , profile of self to precision is convertedmaking use of knowledge of other user which utilizes same data filtering system .

[0010]

[Means to Solve the Problems ]

このため、プロフィールに記述された情報要求と入力された電子化文書とを比較し、情報要求を満たす電子化文書をプロフィールのユーザに配信する情報フィルタリングシステムにおいて、プロフィールごとに配信された電子化文書の情報を参照履歴情報として保存し、あるプロフィールについて、その参照履歴情報が保存された他のプロフィールのうちその参照履歴情報との類似度が所定値よりも高いプロフィールを特定し、特定された他のプロフィールに基づきプロフィールを更新する。

参照履歴情報は、プロフィールに基づき配信された電子化文書の文書識別子と配信された電子化文書に対するユーザの興味との適合を示す適合度を含むようにする。

[0011]

#### 【発明の実施の形態】

以下、本発明の一実施例を図面を用いて説明する。

[0012]

図1に、本発明の情報フィルタリングシステムの実施例の全体構成を示す。

情報フィルタリングシステムは、インターネットまたはイントラネットに接続されている。

システムは、インターネット/イントラネットを介して情報プロバイダ1から送信された情報(WWW文書等)をフィルタリングし、各ユーザに配送する。

情報プロバイダ1は、情報を送信するサービスの提供者であり、ネットワーク上には複数のプロバイダが存在する。

情報フィルタリングシステムには、異なるリソースからの情報、例えば、複数のプロバイダから送信された情報やインターネット/イントラネットなどからInternet Robotと呼ばれるようなエージェント(代理人)システムを用いて収集された情報が配信される。

[0013]

CPU2は、メモリや各種ファイル上にあるプログラムやデータを用いて情報フィルタリング処理を行う中央処理装置である。

メモリ3には、各種の処理を行うプログラムやデータがロードされる。

プログラムは、図示しない例えば磁気記録媒体、光記録媒体のようなコンピュータ読取可能な記録媒体に格納されてシステムに供給され、

Because of this , information need which is described to profile to compare electronic document which is inputted , to retain data of electronic document which is transmitted in every profile in data filtering system which transmits electronic document which fills up the information need to user of profile , as reference history data concerning ascertain profile , similarity of reference history data among other profile where reference history data is retained high other profile specific is done in comparison with specified value , profile is renewed on basis of the other profile which specific is done.

Try reference history data , to include fidelity conformity of interest of user for electronic document which document identifier of electronic document which is transmitted on basis of profile is transmitted is shown.

[0011]

#### [Embodiment of the Invention]

Below, one Working Example of this invention is explained making use of drawing .

[0012]

In Figure 1 , entire constitution of Working Example of data filtering system of this invention is shown.

data filtering system is connected to Internet or intranet .

Through [intaa ]/intranet , data (WWWdocument etc) which was transmitted from the data provider 1 filtering it does system , delivers in each user .

As for data provider 1 , with provider of service which transmits the data , provider of plural can exist on network .

data which was collected making use of agent (representative ) kind of system which is called internet Robot from data and [intaa ]/intranet etc which were transmitted from provider of data , for example plural from different resource is transmitted in data filtering system .

[0013]

CPU 2 is central processing unit which treats data filtering making use of program and data which is on memory and various file .

program and data which treat various load are done in memory 3 .

program being housed in computer readable recording medium like unshown for example magnetic recording medium , optical recording medium , is supplied by the

メモリ 3 にロードされる。

また、システムは、プロフィールファイル 4、参照履歴ファイル 5、類似ユーザ情報ファイル 6、アドレス情報ファイル 7、文書ファイル 8 を有する。

[0014]

プロフィールファイル 4 には、ユーザごとのプロフィールが格納されている。

プロフィールの一例を図 2 に示す。

この例では、プロフィールとして単語を設定する。

図 2 の例に示されるプロフィールを持つユーザは、暗号関係の情報に興味があり、暗号に関連する語として、「暗号」、「公開鍵」、「PGP」、「セキュリティ」という語を設定している。

[0015]

参照履歴ファイル 5 は、ユーザごとに、ユーザが過去に参照した文書について適合度を判定した結果を、参照履歴として格納する。

各ユーザの参照履歴の一例を図 3 に示す。

参照履歴は、各ユーザに配送された文書 ID とその適合度との組の集合である。

ユーザ番号は、情報フィルタリングシステムのユーザの ID 番号であり、各ユーザに対して配信された文書 ID が参照文書 ID として登録される。

適合度は、ユーザが入力した適合度を数値化したものであり、「1」が適合している(ユーザの興味と合致している)場合を、「0」が適合していない(ユーザの興味と合致していない)場合を示す。

[0016]

類似ユーザ情報ファイル 6 は、各ユーザごとに、興味の種類を類似するユーザ(「類似ユーザ」という)の情報を格納する。

各ユーザの参照履歴が類似しているユーザはその興味の種類を類似するとみなして、その ID を類似ユーザとして登録する。

図 4 の例では、ユーザ #1 の類似ユーザとして、ユーザ #2、#7 が登録されている。

[0017]

アドレス情報ファイル 7 は、協調して動作する他のフィルタリングシステムの所在を例えば IP-address などの形で格納する。

system, load is done in memory 3.

In addition, system has profile file 4, reference history file 5, similar user data file 6, address data file 7, document file 8.

[0014]

profile every of user is housed in profile file 4.

one example of profile is shown in Figure 2.

With this example, term is set as profile.

user which has profile which is shown as example of Figure 2 is a interest in data related to code, "Code", "public-key", "PGP", "security" with sets the word which is said as word which it is related to code.

[0015]

In every user, it houses reference history file 5, result of deciding fidelity concerning document which user refers to in subtracting, as reference history.

one example of reference history of each user is shown in Figure 3.

Reference history is gathering of group of document ID and the fidelity it is delivered in each user.

user number with ID number of user of data filtering system, document ID which is transmitted vis-a-vis each user is registered as reference document ID.

Fidelity being something which fidelity user inputs digitization is done, in (Coinciding with interest of user, it is) case where "1" has conformed, shows (It does not coincide with interest of user.) case where "0" has not conformed.

[0016]

Similar user data file 6 houses data of user ("Similar user" With you call) where interest resembles each every user.

user to which reference history of each user resembles regarding that interest resembles, registers ID as similar user.

With example of Figure 4, user #2, #7 is registered as similar user of user #1.

[0017]

address data file 7, cooperating, houses location of other filtering system which operates in for example IP-addressor other shape.

例えば、大きな企業などでは、物理的に離れた個所に事業所/営業所などが存在する場合がある。

このような場合に、フィルタリングシステムをそれぞれの事業所(営業所)に設け、各フィルタリングシステム同士で所有するデータのやり取りが行えれば便利である。

そのため、対象となるシステムの所在をアドレス情報ファイルに記述しておく。

[0018]

文書ファイル 8 は、各ユーザごとに情報配信用に各ユーザのみが参照可能なファイルを有する。

ユーザは情報フィルタリングシステムに接続された端末から、それぞれのファイルにアクセスすることにより、フィルタリングされた文書の配信を受ける。

[0019]

メモリ 3 に格納されるプログラムについて説明する。

入力情報マネージャプログラム 31 は、情報プロバイダ、インター/イントラネットから得られる情報を管理する。

例えば、複数のリソースから同時に多くの情報を受信した場合などに、計算機に対する負荷が大きくなり過ぎるのを防止するため、図示しないファイルに情報を一旦バッファリングすることにより、情報フィルタリングシステムの負荷を調整する。

また、様々なリソースから得られた情報を一括して管理し、配信された文書に対して文書 ID を付与する。

[0020]

情報フィルタプログラム 32 は、各ユーザごとに設定されたプロフィールと配送された文書との比較を行い、プロフィールに合致する程度が高い文書を配信すべき文書として指定する。

[0021]

適合度入力プログラム 33 は、配信された文書に対し、各ユーザにより適合度の入力を受ける。

適合度は、配信された文書が各ユーザの要求と合致している程度を表わす。

適合度は、ユーザの要求に合致するか否かの

With for example large industry etc, there are times when operations center /sales office etc exists in site which leaves to physical .

When it is a this way, it provides filtering system in respective operations center (sales office ) and if it can exchange data which is owned with each filtering system , it is convenient.

Because of that, location of system which becomes object is described to address data file .

[0018]

As for document file 8, in each every user each only user has there reference possible file in one for data transmission.

user receives transmission of document which filtering is done from terminal which is connected to data filtering system , by access doing in respective file .

[0019]

You explain concerning program which is housed in memory 3.

input information manager program 31 manages data which is acquired from data provider , [intaa /intranet .

When many data are received simultaneously from resource of the for example plural etc, in order to prevent fact that load for computer becomes too large, load of data filtering system is adjusted by buffering doing data once in unshown file .

In addition, lumping together data which is acquired from the various resource , it manages, it grants document ID vis-a-vis document which is transmitted.

[0020]

data filter program 32 compares with profile which is set each every user and document which is delivered, appoints as document which should transmit document where extent of coinciding to profile is high.

[0021]

Fidelity input program 33 receives input fidelity vis-a-vis the document which is transmitted, with each user .

Fidelity displays extent document which is transmitted coinciding with demand for each user .

Fidelity inputs with two alternative of whether or not which

二者択一で入力する。

5段階評価などで入力することも可能である。

以下の例では、yes/no の二者択一で入力する (図 6 参照)。

[0022]

類似度判定プログラム 34 は、各ユーザの参照履歴を比較し、類似する参照履歴を持つユーザを発見する。

[0023]

プロフィール更新プログラム 35 は、類似度判定プログラム 34 によって抽出された類似ユーザのプロフィールを参照して、各ユーザのプロフィールを更新する。

[0024]

暗号化・復号プログラム 36 は、情報をネットワークを介して送信する際、情報の暗号化/復号化を行う。

[0025]

以下、本実施例での情報フィルタリングシステムにおける各処理を説明する。

[0026]

(1) 文書の配信・表示処理 (図 5)

情報フィルタリングシステムは、文書が配送されているかどうかをチェックする (ステップ 11)。

情報フィルタプログラム 32 を起動し、プロフィールファイル 4 に格納された各ユーザのプロフィールと配送された文書とを比較し、配信するユーザを決定する (ステップ 12)。

ステップ 12 は次のような手順で行われる。

まず、配送された文書を形態素解析し、タームに分割し、助詞などの機能語を削除して配送文書に含まれるターム集合 IT を作成する。

次に、プロフィールファイル 4 から、各ユーザのプロフィールを逐次取り出し、各プロフィール中に格納されているターム集合 PT (以下、「プロフィールターム」と) IT を比較することによって、プロフィールと配送された文書との一致度 M を計算する。

[0027]

$$M = |PT \cap IT| / |PT \cup IT|$$

M の値が予め定められた閾値より大きいユー

coincides to demand for user .

Also it is possible to input with 5 -step ranking etc.

With example below, it inputs with two alternative of yes/no (Figure 6 reference).

[0022]

similarity decision program 34 compares reference history of each user, discovers user which has reference history which resembles.

[0023]

profile renewal program 35 referring to profile of similar user which is extracted with similarity decision program 34, renews profile of each user .

[0024]

encryption /decoding program 36 when data through network , transmitting, does the encryption /decoding of data .

[0025]

Below, each treatment in data filtering system with this working example is explained.

[0026]

Transmission & display process of (1) document (Figure 5 )

check it does data filtering system whether or not document is delivered, (step 11 ) .

data filter program 32 is started, profile of each user which is housed in the profile file 4 document which is delivered is compared, user which is transmitted is decided (step 12 ) .

step 12 following way is done with protocol .

First, morphological analysis it does document which is delivered, divides into [taamu ], deletes particle or other function word and it draws up [taamu ] gathering IT which is included in delivery document .

Next, from profile file 4, profile of each user is removed one by one, [taamu ] gathering PT which is housed in each profile (Below, " [purofairutamu ] ") with by the fact that IT is compared, degree of agreement M of the profile and document which is delivered is calculated.

[0027]

$$M = |PT \cap IT| / |PT \cup IT|$$

user which is larger than threshold value where it can decide

ザを配信の対象とする。

【0028】

なお、一致度 M の計算式は上記の数式に限定されない。

また、比較対照とするタームは、語彙の特殊性または出現の頻度などによって限定することも可能である。

【0029】

配信先と判定されたユーザに対し文書を配信する(ステップ 13)。

本実施例では、文書ファイル 8 中の配信先ユーザのファイルにその文書を格納する。

情報の配信方法としては、任意の方式が採用できる。

【0030】

各ユーザから、配送された文書を表示する要求があったかどうか調べる(ステップ 14)。

表示要求がなければ、ステップ 11 に戻って待機する。

表示要求を受けると、文書ファイル 8 中の配信文書の表示を要求したユーザのファイルから、配信された文書を取り出し、端末に表示する処理を行う(ステップ 15)。

ユーザに表示される端末画面の例を図 6 に示す。

文書ファイル 8 に格納された配信文書には、最新(もしくは未読)を示すフラグをたてておき、表示要求を受けると最新の(もしくは、未読の)情報を表示する。

各ユーザは配信文書に適合度を入力する。(ステップ 17)。

例えば、図 6 において、記事ごとに配置された「適合」ボタンによって適合度を入力する。

配信文書がユーザの興味に適合していれば「適合」ボタンを押下し、そうでなければ何もしない。

このように入力された適合度は数値化され、配信文書 ID とともにメモリの所定のテーブル(図示せず)に一時的に保存される(ステップ 18)。

配信文書が複数存在する場合には、この処理を繰り返す。

value of M beforehand is designated as object of transmission.

【0028】

Furthermore, computational formula of degree of agreement M is not limited in above-mentioned mathematical formula.

In addition, [taamu] which is made comparative reference limiting with special characteristic of lexicon or frequency etc of appearance is possible.

【0029】

As ahead transmitting document is transmitted vis-a-vis user which is decided (step 13).

With this working example, ahead transmitting in document file 8 document is housed in file of user.

As distribution method of data, you can adopt system of option.

【0030】

You inspect, whether or not from each user, there was request which indicates document which is delivered, (step 14).

display request not becoming, returning to step 11, you wait.

When display request is received, document which from file of the user which requires indication of distribution Nobufumi book in document file 8, is transmitted is removed, treatment which is indicated in terminal is done, (step 15).

Example of terminal screen which is indicated in user is shown in the Figure 6.

When, flag which shows recent (Or not yet reading) is built in distribution Nobufumi book which is housed in document file 8, display request is received, (Or, not yet reading) data of recent is indicated.

Each user inputs fidelity into distribution Nobufumi book. (step 17).

In for example Figure 6, it inputs fidelity with "Conformity" button which is arranged in every article.

If distribution Nobufumi book has conformed to interest of the user, it does to push down "Conformity" button, so does not become at all does not do.

this way fidelity it is inputted is done, digitization with the distribution Nobufumi book ID in predetermined table (not shown) of memory is retained temporary (step 18).

When distribution Nobufumi book plural it exists, this treatment is repeated.



ユーザが配信文書の表示を終了すると、一時保存されている配信文書 ID と適合度とを参照履歴ファイル 5 に格納する(ステップ 19)。

[0031]

本実施例では、参照履歴はプロフィールに対応して格納されている。

プロフィールはユーザの特定の興味に対応して設定され、その興味は経時的に変化しないものであるという前提に立つ。

プロフィールはその特定の興味をよりよく反映するように洗練されていく。

したがって、参照履歴から類似ユーザを発見するためにはデータ量が多いほど望ましいところから、参照履歴はプロフィールの精密化処理の実行にかかわらず、継続的に蓄積されていくことが望ましい。

しかし、参照履歴を格納しているハードディスクなどの記憶装置の容量には限りがあるため、ファイルにデータを格納しようとした際にあふれ(容量不足)が発生する場合がある。

以下では、このような場合に対処するための代表的な方法について図 7 の処理フローを用いて説明する。

[0032]

参照履歴を格納する際、ディスクのあふれがないかどうか調べる(ステップ 191)。

あふれがなければ、参照履歴を格納する(ステップ 192)。

あふれが検出された場合には、圧縮処理を行う。

圧縮処理の対象となるユーザを決定する(ステップ 193)。

圧縮処理対象とするユーザは、大きな参照履歴を持つユーザまたは使用頻度が低いユーザを対象とできる。

圧縮処理対象とするユーザの参照履歴を圧縮する(ステップ 194)。

数回圧縮を繰り返しても、あふれが発生する場合には処理を停止し、管理者に連絡するなどの対策を行う。

[0033]

参照履歴の圧縮は、参照履歴中の優先度が低い文書 ID の情報を削除する。

When user ends indication of distribution Nobufumi book, the distribution Nobufumi book ID and fidelity it is retained at one time are housed in reference history file 5 (step 19).

[0031]

With this working example, as for reference history corresponding to profile, it is housed.

profile is set corresponding to specific interest of user, the interest stands in premise that is something which does not change in timewise.

profile is refined in order to reflect specific interest well.

Therefore, in order to discover similar user from reference history, when data amount is many, from desirable place, as for reference history compilation it is desirable in continual regardless of execution of precision conversion treatment of profile, to be done.

But, because there is limit in capacity of hard disk or other storage device which houses reference history, it overflows occasion where it tries to house the data in file there are times when (deficient capacity) occurs.

At below, when it is a this way, you explain making use of process flow of Figure 7 concerning representative method in order to cope.

[0032]

When housing reference history, you inspect, whether or not there is not overflow of disk, (step 191).

Overflow not becoming, reference history is housed (step 192).

When overflow is detected, compression is done.

user which becomes object of compression is decided (step 193).

user which is made compression object can designate user where the user or use frequency which has large reference history is low as the object.

Reference history of user which is made compression object is compressed (step 194).

several times compression over again, when overflow occurs, treatment is stopped, or other fix which you communicate to administrator is done.

[0033]

Compression of reference history deletes data of document ID where priority in reference history is low.

各文書 ID の優先度は、日時(新しいものほど優先度を高くする)によって判定する。

あるいは、その文書がどの程度、情報フィルタリングシステムのユーザに配信されたか(配信されたユーザの多いものほど優先度を高くする)によって判定する。

これは、参照履歴が類似ユーザの判定に用いられるものであり、その文書を参照した他のユーザが少ないければその判定にほとんど影響しないためである。

【0034】

なお、参照履歴ファイル 5 の全体でなく、各ユーザごとの参照履歴ファイルに上限を設けておき、上限に達した時点で、参照履歴ファイル圧縮を行うようにしてもよい。

この方法では、参照履歴ファイル全体の大きさの上限が定められる。

【0035】

(2)プロファイルの精密化処理(図 8)プロファイル精密化処理を起動する(ステップ 21)。

タイマーなどにより定期的に起動してもよく、前回の精密化処理後に蓄積された参照履歴が一定以上の大きさになったことを検知して起動してもよい。

また、類似ユーザ情報ファイル 6 を参照し、類似ユーザのプロファイルについて変更があった場合に、処理を起動することでもできる。

【0036】

プロファイルの精密化はあらかじめ設定されているため、類似ユーザを探索する処理はフィルタリング処理とは独立して行うことができる。

その意味で、一般の情報検索と異なり、検索条件の精密化にはリアルタイム性を要しない。

そのため、類似ユーザの探索処理には計算負荷が大きいが、実用上の問題はない。

また、ユーザによる精密化起動指示を待たず、上述のような条件により処理起動もしくはメール等により起動要請を通知することが望ましい。

これは、情報フィルタリングの文書ファイル 8 の容量の問題から、フィルタリング処理は文書が配送される度に行うことができる。

この場合、長期にわたってプロファイルの更新を実行されないことにより、ユーザの興味に合

It decides priority of each document ID, with day and time (About new ones priority high it does).

Or, document which extent, was transmitted in user of data filtering system, it decides with (About those where user which is transmitted is many priority high it does).

As for this, being something where it can use for decision of the similar user reference history, other user which refers to the document is little, \* is because it does not have an influence on that decision for most part.

【0034】

Furthermore, not to be an entirety of reference history file 5, it provides the upper limit in reference history file of each every user, with time point which reaches to upper limit, it is possible to do reference history file compression.

With this method, it can decide upper limit of size of reference history file entirety.

【0035】

precision conversion treatment (Figure 8) profile precision conversion treatment of (2) profile is started (step 21).

periodically it is possible to start with timer etc, detecting the fact that reference history which compilation is done becomes the size above uniformity after precision conversion treating of previous time, to start it is possible.

In addition, you refer to similar user data file 6, when there is modification concerning profile of similar user, you can also start treatment.

【0036】

Can convert profile precision because it is set beforehand, the treatment which searches similar user becoming independent with filtering treatment, to do.

In that sense, real time behavior is not required in precision conversion of the retrieval condition unlike general information retrieval.

Because of that, calculated load is large to search treatment of the similar user, but there is not a problem in regard to utility.

In addition, it does not wait for precision conversion starting display with user, it is desirable to notify starting request with the condition an above-mentioned way with treatment starting or mail etc.

This, from problem of capacity of document file 8 of data filtering, must execute filtering treatment in degree where document is delivered.

In case of this, is because there is a possibility document which coincides to interest of user by renewal of profile

致する文書が配送されなくなるおそれがあるためである。

[0037]

精密化処理を要求する要求ユーザの参照履歴 Rb を参照履歴ファイル 5 から取得する(ステップ 22)。

また、要求ユーザのプロファイルのプロファイルファイル 4 から取得する(ステップ 23)。

[0038]

その後、他のユーザ(「比較対照ユーザ」という)の参照履歴 Rt を参照履歴ファイル 5 から取得する(ステップ 25)。

ここで、比較対照ユーザは、自情報フィルタリングシステムのみではなく、アドレス情報ファイル 7 に記述された他の情報フィルタリングシステムをも対象とすることが望ましい。

なお、異なるシステム間で情報をやり取りする際には、送信側の計算機の暗号化・復号プログラム 36 によって暗号化を行った後で送信し、受信側の計算機の暗号化・復号プログラム 36 によって復号する。

[0039]

要求ユーザの参照履歴 Rb と比較対象ユーザの参照履歴 Rt とを比較し、類似度を決定する(ステップ 26)。

比較は、次のように行う。

参照履歴 Rb および Rt から、適合度が「1」の文書 ID 集合である適合文書履歴 Rb'、Rt' を抽出する。

このとき、類似度 sim を次のように定める。

[0040]

s	m	Rb	n	R				Rb	U	R		
~~~~~	~~~~~	~~~~~	~~~~~	~~~~~				~~~~~	~~~~~	~~~~~		
s	m	Rb	X	R				Rb	X	R		

義しているが、情報フィルタリングシステムに入力される情報の量が少ない場合には、類似度をうまく設定できない場合がある。

別の類似度を求めるアプローチとして、参照履歴中の文書間の類似性を求め、類似文書に関する適合度を用いることで類似度を定義することも可能である。

類似文書の判定は、例えば、文書中の単語の出現頻度を用いて各文書の特徴付ける単語の

nothing executed over long period, stopping being delivered.

[0037]

Reference history Rb of required user which requires precision conversion treatment is acquired from reference history file 5 (step 22).

In addition, profile of required user is acquired from the profile file 4 (step 23).

[0038]

After that, reference history Rt of other user ("comparative reference user" With you call) is acquired from reference history file 5 (step 25).

Here, comparative reference user is not only our data filtering system, it is desirable to designate also other data filtering system which is described to address data file 7 as the object.

Furthermore, when exchanging data between different system, after, doing encryption with encryption \*decoding program 36 of computer of transmitting side it transmits, the decoding does with encryption \*decoding program 36 of computer of called side.

[0039]

Reference history Rb of required user and reference history Rt of comparative object user are compared, similarity is decided (step 26).

It compares, following way.

From reference history Rb and Rt, conformity document history Rb\*, Rt\* where fidelity is document ID gathering of "1" is extracted.

At time of this, similarity sim is decided following way.

[0040]

Artificial limb \* it is, but when quantity of data which is inputted into data filtering system is small, there are times when similarity cannot be set well.

Also it is possible to define similarity by fact that it seeks the similarity between document in reference history as approach which seeks another similarity, it uses fidelity regarding similar document.

Decision of similar document extracts gathering (Feature word gathering) of term which characterizes each document,

集合(特徴語集合)を抽出し、特徴語集合同士を比較することによって行える。

2個の文書D1、D2から特徴語集合F1、F2を抽出する。

特徴語集合は、各文書から形態素解析処理によって単語を抽出し、名詞、動詞などの内容語の出現頻度をカウントし、予め定められた閾値以上の頻度を持つ内容語を抽出することによって得られる。

特徴語集合の文書類似度Fsimは例えば、以下の式を用いる。

[0041]

$$Fsim = |F1 \cap F2| / |F1 \cup F2|$$

文書類似度Fsimが予め定められた閾値以上の場合、文書D1とD2とは類似していると判定する。

そして、類似度simの式における $|Rb' \cap Rt'|$ の代わりに、類似する文書数によって類似度simを定義する。

また、類似文書の判定には、他にも、tf-idf(term frequency - inverse document frequency)法を用いたベクトル空間法などを用いることもできる。

他の例については、「J. Mostafa et al., A multilevel Approach to Intelligent Information Filtering: Model, System, and Evaluation, ACM Transactions on Information Systems, Vol. 15, No.4, pp.368-399, October 1997.」などに開示されているため詳細な説明は省略する。

[0042]

参照履歴の類似度simが予め定められた閾値と比較し(ステップ27)、類似度simの大きい比較対象ユーザは要求ユーザと類似した興味を持っていると判定し、対応するプロフィールをプロフィールファイル4から取得し、比較対象ユーザのプロフィールから要求ユーザのプロフィールに含まれないタームを抽出して、追加タムリスト(図示せず)に格納する(ステップ28)。

全ての他のユーザについて以上の処理が終了すれば(ステップ24)、得られた追加タムリスト中のタームを要求ユーザのプロフィールに追加し、新規プロフィールとしてプロフィールファイルに格納する(ステップ29)。

[0043]

making use of extraction frequency of term in the for example document can do by fact that feature word gathering is compared.

Feature word gathering F1, F2 is extracted from document D1, D2 of 2.

Feature word gathering from each document extracts term with the morpheme analysis, counts extraction frequency of noun, verb or other content word, is acquired by fact that the content word which has frequency of threshold or higher which is decided beforehand is extracted.

document similarity Fsim of feature word gathering uses formula below for example.

[0041]

$$Fsim = |F1 * F2| / |F1 \cup F2|$$

When it is a threshold or higher where it can decide document similarity Fsim beforehand, it decides that it resembles document D1 and D2.

And, similarity sim is defined with number of documents which resembles in place of  $|Rb' \cap Rt'|$  in formula of similarity sim.

In addition, to it can decide similar document, to in addition to, also the fact that vector space method etc which uses tf-idf (term frequency - inverse document frequency) method is used.

Concerning other example, because it is disclosed in "J. Mostafa et al., A multilevel Approach to intelligent Information Filtering: Model, system, and Evaluation, ACM transactions on information Systems, Vol.15, No.4, pp.368-399, October 1997." etc, it abbreviates detailed description.

[0042]

By comparison with threshold value where it can decide similarity sim of reference history beforehand (step 27), as for comparative object user where similarity sim is large that it has interest which resembles required user it decides, it acquires profile which corresponds from profile file 4, extracting [taamu] which from profile of comparative object user is not included in profile of required user, it houses in additional [taamurisuto] (not shown) (step 28).

If treatment above ends concerning all other user, (step 24), it adds [taamu] in additional [taamurisuto] which is acquired to profile of therequired user, it houses in profile file as novel profile (step 29).

[0043]

なお、プロフィールとしてタームの論理式(検索式)が与えられている場合には次のようにする。

第一の方法としては、プロフィールの論理式からタームのみを切り出し、タームの集合について同じ処理を行う。

追加されるタームは、今までの検索式に、OR(論理和)で追加する。

第二の方法としては、ステップ 28 で、追加する検索式として各ユーザの検索式を全て抽出し、ステップ 29 で、類似ユーザのプロファイルの検索式を OR で追加する。

この場合、例えば、自分のプロフィールが「暗号 AND 認証」であり、類似ユーザのプロファイルが「暗号 AND セキュリティ」であれば、新しいプロフィールは、「(暗号 AND 認証)OR(暗号 AND セキュリティ)」となる。

得られた検索式を論理式の簡略化方式に従って簡略化してもよい。

[0044]

プロフィールを精密化にあたっては、追加タームリストをユーザに表示し、実際にプロフィールに追加するかどうかをユーザに判断させるような構成とすることも可能である。

その場合、ユーザの判断を支援するための情報を提示することが望ましい。

以下では、ステップ 28 におけるインタラクティブな追加ターム決定方法について図 9 を用いて説明する。

[0045]

追加タームリスト中のタームを表示する(ステップ 281)。

表示画面の例を図 10 に示す。

図 10 では、追加タームとして、「認証」、「電子署名」、「楕円関数」が表示されている。

ユーザは、これらの追加タームを参照し、それぞれのタームについて自分のプロフィールに追加するかどうかを「yes」/「no」で判断し、追加ボタンによってシステムに追加の可否を指示する。

[0046]

ユーザは、判断の参考になる情報を望む場合は、支援情報ボタンを押下する。

支援情報ボタンが押下された場合には、追加ターム判定のための支援情報を表示する(ステップ 284)。

Furthermore, when boolean (search statement) of [taamu] is given as profile it makes following way.

As method of first, it cuts only [taamu] from boolean of the profile, it treats similarly concerning gathering of [taamu].

To now it adds [taamu] which is added, to search statement, with OR (disjunction).

As second method, with step 28, search statement of each user all is extracted as search statement which is added, with step 29, search statement of profile of similar user is added with OR.

In case of this, for example your own profile being "Cryptographic AND certification", if the profile of similar user is "Cryptographic AND security", new profile becomes " (Cryptographic AND certification) OR (Cryptographic AND security )" with.

Following search statement which it acquires to simplification system of boolean, it is possible to simplify.

[0044]

Also it is possible to make kind of configuration which judges profile in user whether or not it indicates additional [taamurisuto] in the user at time of precision conversion, adds to profile actually.

In that case, it is desirable to present data in order to support judgement of user.

At below, concerning [intarakutibu] additional [taamu] determination method in step 28 you explain making use of Figure 9.

[0045]

[taamu] in additional [taamurisuto] is indicated (step 281).

Example of display screen is shown in Figure 10.

With Figure 10, "Certification", "electron signature", "ellipse function" is indicated addition [taamu] as.

"yes" / It judges with "no", with add button display does additional yes or no in system whether or not user refers to these additional [taamu], add to your own profile concerning respective [taamu].

[0046]

user when data which becomes reference of judgement is desired, pushes down supporting information button.

supporting information button pushing when, supporting information for additional [taamu] decision is indicated (step 284).

図 11 に、追加ターム判定支援情報の表示例を示す。

支援情報は、プロフィールに追加タームを加えることによって、新たにどのような文書が入力できるかを例示するものである。

「認証」という追加タームに関し、「認証」を含み、類似ユーザのみに配信された文書#3008 が表示されている。

ユーザは、この文書を読み、「認証」という追加タームを追加するかどうかの適切さを判断する。

この処理は、要求ユーザに配信された文書 ID リストと類似ユーザに配信された文書 ID リストの差分を抽出し、追加タームのそれぞれについて抽出された文書を探索し、各タームごとに各タームを含む文書 ID リストを表示することにより実現できる。

【0047】

また、他の支援情報として、類似ユーザのユーザ情報を表示してもよい。

ステップ 27(図 8)において類似度が高いと判定されたユーザのユーザ情報を取得する。

ユーザ情報とは、例えば、ユーザの氏名、所属部署、業務内容、電子メールアドレス、電話番号、住所などである。

そして、このユーザ情報を追加タームリストと共に表示する。

この場合、各ユーザは、支援情報として表示可能なユーザ情報に制限を加える。

例えば、社内などで秘密のプロジェクトなどに参加しているユーザなどの場合である。

また、インターネット上で広く利用されるシステムなどにおいても、プライバシー上の問題がある。

そこで、各ユーザは他のユーザに対して開示してもよい情報の範囲を予め定めておく。

一切フィルタリング情報を開示しない場合は、プロフィールに対して、プロフィール精密化処理の対象としても良いかどうかを示すフラグを立てておき、このフラグが立てられているプロフィールに対応する参照履歴は、ステップ 25 の参照履歴取得処理の対象外とする。

他にも、プロフィールを参照することは可能だが、ユーザの氏名などは秘密にするなどのバリエーションが考えられる。

In Figure 11, display example of additional [taamu] decision supporting information is shown.

supporting information is something which is illustrated by fact that theadditional [taamu] is added to profile, it can procure which kind of document anew.

"Certification" With it regards additional [taamu] which is said, including "Certification",document #3008 where is transmitted in only similar user is indicated.

Appropriateness whether or not user reads this document, "Certification" withadds additional [taamu] which is said of, it judges.

this treatment can extract difference of document IDlist which is transmittedin document IDlist and similar user which are transmitted inrequired user, can search document which is extracted additional[taamu] concerning respectively can actualize by indicating document IDlist whichincludes each [taamu] each every [taamu].

[0047]

In addition, it is possible to indicate user data of similar user as other supporting information.

user data of user which is decided that similarity is high in the step 27 (Figure 8) is acquired.

user data, is name, affiliation post and business content, electronic mail address, telephone number, address etc of for example user.

And, this user data is indicated with additional [taamurisuto].

In case of this, each user adds restriction to displayable user data as the supporting information.

In case of user or other which participates in project etc of secretwith such as for example intracompany is.

In addition, regarding system etc which is widely utilized on the Internet, there is a problem on privacy.

Then, each user decides range of data which it is possible to disclose vis-a-vis other user beforehand.

Altogether, when filtering data is not disclosed, vis-a-vis profile, as object of profile precision conversion treatment flag which shows whether or not it is good, is raised, reference history which corresponds to profile where this flag is raised makes outside object ofreference history acquisition treatment of step 25.

It is possible to in addition to to refer to profile, but as for the name etc of user you can think or other variation which is made secret.

[0048]

追加ボタンが押下されれば(ステップ 282)、ユーザが選択した追加タームが残るように追加タームリストを変更する(ステップ 285)。

[0049]

なお、ステップ 26(図 8)の参照履歴間の類似度  $s_{im}$  の判定処理において、比較対象ユーザの参照履歴として適合度が「1」の文書のみ(適合文書履歴  $Rt'$ )を対象とした。

これに対して、配信された全ての文書を含む  $Rt$  を用いることも可能である。

比較対象ユーザの参照履歴として、 $Rt$  と  $Rt'$  を用いる場合では、以下のような違いがある。

[0050]

$Rt$  は、比較対象ユーザが設定しているプロフィールによって配信される情報全てであり、 $Rt$  とプロフィールとは一対一に対応している。

よって、 $Rb'$  (要求ユーザの適合文書履歴)と  $Rt$  が類似していれば、 $Rt$  に対応するプロフィール中のタームは、要求ユーザにとって追加すべきタームであるといえる。

しかし、 $Rt$  中に比較対象ユーザにとっても興味に合わない文書が含まれている可能性があるため、これらのノイズが類似ユーザを探す場合に精度を下げる可能性がある。

[0051]

一方、 $Rt'$  は、比較対象ユーザが興味を持つ文書のみが含まれている。

よって、 $Rb'$  と  $Rt'$  とが類似していれば、比較対象ユーザの興味と要求ユーザの興味とは一致している可能性が非常に高い。

しかし、比較対象ユーザが興味をもたなかった文書に対応するタームが追加される可能性が生じることで精度を下げる可能性がある。

そのため、いずれの参照履歴( $Rt$ または  $Rt'$ )を用いるかは、目的に応じて決定すればよい。

[0052]

配信文書に対するユーザによる適合度の入力の手間を軽減する例を説明する。

[0053]

第一の方法としては、ステップ 26(図 8)の類似度判定処理において、配信された文書の文書 ID

[0048]

If pushing add button, (step 282), in order for additional [taamu] which the user selects to remain, additional [taamurisuto] is modified (step 285).

[0049]

Furthermore, fidelity only document of "1" designated (Conformity document history  $Rt'$ ) as object in decision process of similarity  $s_{im}$  between reference history of the step 26 (Figure 8), as reference history of comparative reference user.

Vis-a-vis this, also it is possible to use  $Rt$  which includes all document which is transmitted.

As reference history of comparative object user, with when  $Rt$  and  $Rt'$  are used, like below there is difference.

[0050]

$Rt$ , with data all which is transmitted with profile which the comparative reference user has set,  $Rt$  and profile corresponds to one-to-one.

Depending, if  $Rb'$  (Conformity document history of required user) with  $Rt$  resembles, you can call the [taamu] in profile which corresponds with  $Rt$ , that it is a [taamu] which it should add for required user.

But, because there is a possibility where document which is not agreeable to interest for comparative object user is included in  $Rt$ , when these noise search similar user, there is a possibility which lowers precision.

[0051]

On one hand, as for  $Rt'$ , only document where comparative object user has the interest is included.

Depending, if  $Rb'$  and  $Rt'$  resemble, interest of comparative object user and interest of required user possibility which agrees extremely high.

But, there is a possibility which lowers precision by fact that the possibility where [taamu] which corresponds to document where comparative reference user did not have interest is added occurs.

Because of that, if it should have decided according to objective it uses no reference history ( $Rt$  or  $Rt'$ ).

[0052]

Example which lightens labor of input fidelity with user for distribution Nobufumi book is explained.

[0053]

As method of first, it is possible to decide similarity between user with document ID of document which is transmitted in

により、ユーザ間の類似度を決定することが可能である。

すなわち、 $\text{sim} = |\text{Rb} \cap \text{Rt}| / |\text{Rb} \cup \text{Rt}|$  として算出する。

この場合、要求ユーザの適合文書履歴を用いる場合と比較すると精度が若干低下する可能性があるが、追加タームをチェックすることにより、適切なタームをプロフィールに追加するようにする。

【0054】

第二の方法としては、配信文書の表示と適合度の入力とを連動させることにより、明示的な適合度の入力をなくす。

この場合の表示方法の例を図 12 に示す。

図 12 は、配信文書のタイトル一覧を示している。

タイトルがなければ先頭から予め定められた文字数のみを表示する。

あるいは、「田中他、朝倉日本語新講座運用 2・人文系研究のための言語データ処理入門、朝倉書店(1983)」などに開示されているような、重要語や重要文を抽出する技術によって抽出される重要語や重要文を記事と対応して表示しても良い。

ユーザは、ステップ 15(図 5)において、記事の一覧を参照して、自分が読みたいと判断した文書を選択し、選択した記事に対応する「表示」ボタンを押下することにより、記事の内容を表示させることができる。

ここで、ユーザが選択した文書は適合と、選択されなかった文書は不適合とみなす。

【0055】

第三の方法としては、類似ユーザの適合度情報を利用し、ユーザが適合/不適合の判定をしなかった文書に対しても、類似ユーザの判定結果を援用して適合度情報を与えることも可能である。

ユーザ A の類似ユーザとしてユーザ B が類似ユーザ情報ファイル 6 に登録されているとする。

ユーザ A が別途、ユーザ B を指定しても良い。

配信文書中、ユーザ A が適合度を指定しなかった文書がユーザ B の参照履歴中に存在するか

similarity decision process of step 26 (Figure 8).

It calculates as namely,  $\text{sim} = |\text{Rb} * \text{Rt}| / |\text{Rb} * \text{Rt}|$ .

In case of this, when it compares with case where the conformity document history of required user is used there is a possibility where precision decreases somewhat, but try to add appropriate [taamu] to profile by check doing additional [taamu].

【0054】

As second method, by indication of distribution Nobufumi book and input fidelity linkage doing, input explicit fidelity losing.

Example of display method in this case is shown in Figure 12.

Figure 12 has shown title list of distribution Nobufumi book.

title not becoming, only number of characters which is beforehand decided from head indicates.

Or, it seems that is disclosed in "Tanaka other things, Asakura Japanese new chaired laboratory use 2 \* language data processing introduction, Asakura Shoten for humanity typeresearch (1983)" etc, corresponding with the article, it is good indicating important word and important sentence which are extracted with technology which extracts important word and the important sentence.

user referring to list of article in step 15 (Figure 5), can select document which it judges, that by his want to read, can indicate content of article by pushing down "Indication" button which corresponds to article which it selects.

Here, as for document which user selects as for document which it does not conform and does not select and you regard non-conformity.

【0055】

As third method, making use of fidelity data of similar user, user conformity / vis-a-vis document which does not decide non-conformity, invoking determination result of similar user, also it is possible to give fidelity data.

We assume that user B is registered to similar user data file 6, as the similar user of user A.

user A other way, is good appointing user B.

You inspect whether or not in distribution Nobufumi book, document to which user A does not appoint fidelity exists in



どうかを調べる。

存在した場合には、ユーザ B の参照履歴中の適合度を調べ、ユーザ B による適合度をユーザ A の適合度として援用し、参照履歴に格納する。

【0056】

なお、各ユーザのプロファイルは別々に扱うのではなく、類似する興味を持つユーザを一つのグループとして扱うことも可能である。

類似する興味をもつユーザは、以上で説明した実施例に従って自動抽出しても良いし、人手で設定しても良い。

例えば、同じ業務を担当する複数のユーザは、最初から類似ユーザとして指定することが可能である。

既に類似した興味を持つユーザが分かっている場合には、類似する興味を持つユーザのプロファイルおよび参照履歴をマージしたデータを一時的に作成し、以上で説明した実施例を適用することにより、さらに他の類似した興味を持つユーザを精度良く発見することができ、プロファイルを精密化することが可能となる。

【0057】

さらに、本実施例では、説明を簡単にするため、各ユーザのプロファイルは 1 個であるように説明したが、実際にはユーザが複数の事柄に興味を持つことは多い。

このような場合には、参照履歴をユーザ単位ではなく、プロファイル単位に作成することで、以上で説明した方法が適用可能である。

図 13 に複数のプロファイルを設定した場合の参照履歴の例を示す。

ユーザ #1 が 2 個のプロファイルを持っており、各プロファイルに対して、参照履歴が設定されている。

【0058】

(3) プロファイルの精密化処理(図 14)

次に、プロファイルの精密化処理の別の実施例として、参照履歴を使用しないプロファイル精密化方法について説明する。

本実施例では、プロファイル同士を比較することによってユーザの興味は類似しているかどうかを判定する。

reference history of user B.

When it exists, you inspect fidelity in reference history of the user B, you invoke fidelity with user B house in reference history as fidelity of user A.

【0056】

Furthermore, profile of each user is not to handle separately, also it is possible to handle user which has interest which resembles group of one.

user which has interest which resembles is good doing followingto Working Example which is explained at above, automatic extraction and, it is goodsetting with labor .

As for method of plural which takes charge of for example same business , from first as similar user it is possible to appoint.

When user which has interest which resembles alreadyunderstands, to draw up data which profile and reference history of user which has interest which resembles merge isdone in temporary , furthermore precision to be able discover user which has interest to which other things resemble well by applying Working Example which is explained at above, To precision it becomes possible to convert profile .

【0057】

Furthermore, with this working example , in order to make explanation simple, as lbeen, you explained profile of each user , but actually as for the user having interest in item of plural it is many.

When it is a this way, reference history it is not a user unit , by thefact that it draws up in profile unit , method which is explained atabove is applicable .

Example of reference history when profile of plural is set to Figure 13 is shown.

user #1 has profile of 2, reference history is set vis-a-vis each profile .

【0058】

precision conversion treatment of (3) profile (Figure 14 )

Next, you explain concerning profile precision conversion method which does not use reference history as another Working Example of precision conversion treatment of profile .

With this working example , it decides whether or not interest of user resembles by fact that profile is compared.

この方法の場合、参照履歴を使用しないため、処理が単純であり、ディスクの容量が圧縮できるなどの利点がある。

図 8 に示した第一の例と共通する部分については説明を省略する。

【0059】

要求ユーザのプロファイル P<sub>b</sub> をプロファイルファイル 4 から取得する(ステップ 32)。

比較対象ユーザのプロファイル P<sub>t</sub> をプロファイルファイル 4 から取得する(ステップ 34)。

要求ユーザのプロファイル P<sub>b</sub> と比較対象ユーザのプロファイル P<sub>t</sub> を比較し、類似度 simp を決定する(ステップ 26)。

【0060】

$\text{simp} = |P_b \cap P_t| / |P_b \cup P_t|$  プロファイルの類似度 simp が予め定められた閾値より大きいかどうかを判定し(ステップ 36)、大きい場合には追加タムリストの設定を行う(ステップ 37)。

【0061】

プロファイル同士を直接比較する場合、設定されているタム数が少ないため、類似度を正しく求めることができない場合がある。

これを解決するため、シソーラスを用いる。

シソーラスとは、語と語の意味的關係を記述したデータである。

關係としては、上位概念・下位概念、類義などがある。

シソーラス上で所定の關係を持つ語(例えば、類義語等)を同じ語とみなすことにより、類似度を計算することができる。

【0062】

さらに、ユーザが自分と類似した興味を持つユーザをあらかじめ認識していれば、そのような類似ユーザをユーザが直接指定して、自己のプロファイルを精密化するようにすればよい。

例えば、企業においては、自分と類似または関連する業務を行っている部署や人からそのような類似ユーザを特定できる。

あるいは、実施例で述べたようなプロファイル精密化処理によって、あるユーザが自分と類似した興味を持っていることが分かると同時に、その類似ユーザの所属する部署の業務が自分の興味と関連が深いことが分かったとする。

In case of this method, because reference history is not used, treatment being simple, there is a or other benefit which can compress the capacity of disk.

Explanation is abbreviated example and concerning portion which the common of first which is shown in Figure 8 is done.

【0059】

profile P<sub>b</sub> of required user is acquired from profile file 4 (step 32).

profile P<sub>t</sub> of comparative object user is acquired from profile file 4 (step 34).

profile P<sub>b</sub> of required user and profile P<sub>t</sub> of comparative object user are compared, similarity simp is decided (step 26).

【0060】

It decides and whether or not it is larger than threshold value where it can decide similarity simp of  $\text{simp} = |P_b \cap P_t| / |P_b \cup P_t|$  profile beforehand, (step 36), when it is large, it sets additional [taamurisu to], (step 37).

【0061】

When profile is compared directly, because quantity of [taamu] which is set is small, there are times when it is not possible to seek the similarity correctly.

In order to solve this, thesaurus is used.

thesaurus is data which describes semantic relationship of word and the word.

As relationship, there is overriding concept \*lower level concept, Gitou.

similarity can be calculated by regarding word (for example synonym etc) which has the relationship of predetermined on thesaurus same word.

【0062】

Furthermore, if user which has interest to which user resemble by your beforehand is recognized, user appointing the kind of similar user directly, profile of self to precision to convert should have tried.

Regarding for example industry, by your specific is possible kind of similar user from post and person who do business which it resembles or is related, or.

Or, when in profile precision conversion kind of treatment which is expressed with Working Example, it understands that it has interest to which ascertain user resemble by your we assume that simultaneously, the affiliation of similar user business of post which is done your own interest and relation

このような場合には、その部署に所属する他のユーザのプロファイルも参考にして自分のプロフィールを精密化できる。

この場合には、図8または図14の処理フローにおいて、あらかじめ比較対象ユーザとして指定する単数又は複数のユーザもしくはユーザが属するグループの名称を指定するステップを設ける。

そして、プロフィールの精密化に利用する比較対象ユーザをあらかじめ指定されたユーザもしくはグループに属するユーザに限定してステップ24-29(図8)、ステップ33-39(図14)の処理を行う。

[0063]

#### [発明の効果]

本発明によれば、情報フィルタリングシステムにおいて、自分と類似した興味を持つユーザが設定したプロフィールを利用することにより、自分のプロフィールをより自分の興味に適合したものに修正することが可能となる。

#### [図面の簡単な説明]

[図1]

本発明の実施例である情報フィルタリングシステムのブロック図である。

[図2]

プロフィールの例である。

[図3]

参照履歴の例である。

[図4]

類似ユーザ情報の例である。

[図5]

情報配信・表示処理の処理フローである。

[図6]

配信情報の表示例である。

[図7]

記憶装置のあふれ対策処理の処理フローである。

[図8]

プロフィール精密化処理の処理フローである。

being deep you understood.

When it is a this way, to precision you can convert your own profile in that post profile of other user which affiliation is done reference.

In case of this, step which appoints user of single or multiple which it appoints in process flow of Figure 8 or Figure 14, beforehand as comparative object user or title of group to which user belongs is provided.

And, beforehand are appointed limiting comparative object user which is utilized in precision conversion of profile in user which belongs to user or group which step 24-29 (Figure 8), it treats step 33-39 (Figure 14).

[0063]

#### [Effects of the Invention]

From your own profile you become possible according to this invention, in the data filtering system, by utilizing profile which user which has interest which resemble by your sets, to correct in those which are adapted to your own interest.

#### [Brief Explanation of the Drawing (s)]

[Figure 1]

It is a block diagram of data filtering system which is a Working Example of this invention.

[Figure 2]

It is an example of profile.

[Figure 3]

It is an example of reference history.

[Figure 4]

It is an example of similar user data.

[Figure 5]

It is a process flow of data transmission & display process.

[Figure 6]

It is a display example of transmission data.

[Figure 7]

It is a process flow of overflow fix treatment of storage device.

[Figure 8]

It is a process flow of profile precision conversion treatment.

## 【図9】

インタラクティブな追加ターム決定処理の処理フローである。

## 【図10】

追加タームリストの表示例である。

## 【図11】

追加ターム判定支援情報の表示例である。

## 【図12】

明示的な適合度入力を省略するための配信情報の一覧表示画面の例である。

## 【図13】

各ユーザが複数のプロフィールを設定可能な場合の参照履歴の例である。

## 【図14】

第2の実施例におけるプロフィール精密化処理の処理フローである。

**Drawings**

## 【図1】

[Figure 9]

It is a process flow of [intarakutibu] additional [taamu] decision treatment.

[Figure 10]

It is a display example of additional [taamurisuto].

[Figure 11]

It is a display example of additional [taamu] decision supporting information.

[Figure 12]

It is an example of list display screen of transmission data in order to abbreviate explicit fidelity input.

[Figure 13]

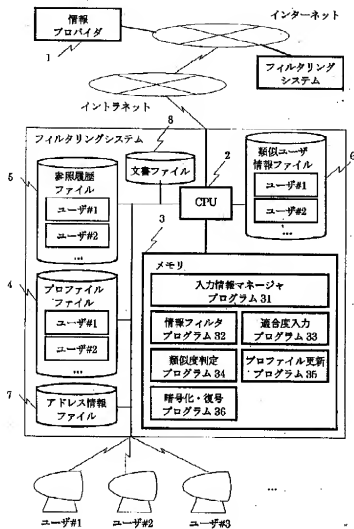
Each user profile of plural settable is example of thereference history when.

[Figure 14]

It is a process flow of profile precision conversion treatment in second Working Example.

[Figure 1]

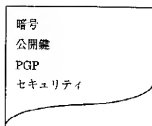
図 1



【図2】

[Figure 2]

図 2



【図3】

[Figure 3]

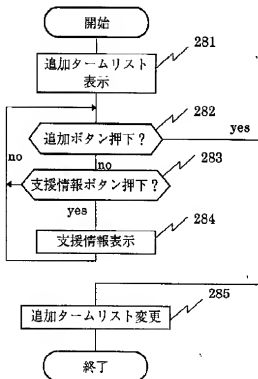
図 3

ユーザ番号					
#1	参照文書 ID	#1024	#1248	#2045	#3042
	適合度	1	0	1	1
#2	参照文書 ID	#1024	#2045	#3042	#3045
	適合度	1	1	1	0
#3	参照文書 ID	...	...	...	...
	適合度	...	...	...	...

【図9】

[Figure 9]

図 9



【図4】

[Figure 4]

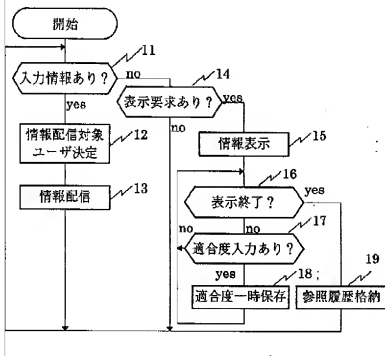
図 4

比較元ユーザ	類似ユーザ		
#1	#2	#7	—
#2	#4	—	—
#3	—	—	—
...	...	...	...

【図5】

[Figure 5]

図 5



【図10】

[Figure 10]

図 10

追加		支援情報	
認証	追加?	yes	no
電子署名	追加?	yes	no
権限関数	追加?	yes	no

【図6】

[Figure 6]

図 6

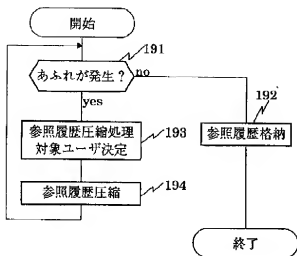
#3043	暗号を利用した...	適合
#4056	セキュリティを...	適合
...		

【図7】

[Figure 7]



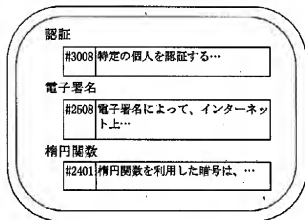
図 7



【図11】

[Figure 11]

図 11



【図13】

[Figure 13]

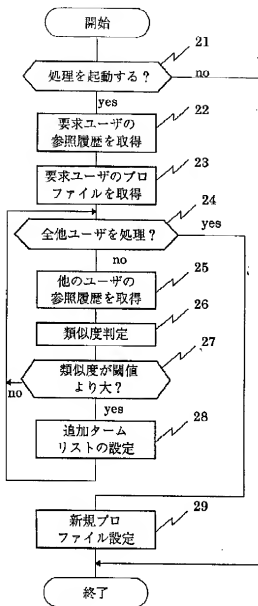
図 13

ユーザ番号	プロフィール番号					
#1	#1	参照文書 ID	#1024	#1248	#2045	#3042
		適合度	1	0	1	1
	#2	参照文書 ID	#1023	#2022	#3022	#3045
		適合度	1	0	1	0
#2	#1	参照文書 ID	...	...	...	...
		適合度	...	...	...	...

【図8】

[Figure 8]

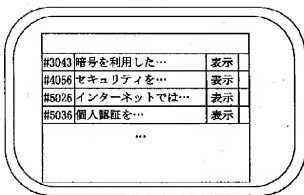
図 8



【図12】

[Figure 12]

図 1 2



#3043	暗号を利用した…	表示
#4056	セキュリティを…	表示
#5026	インターネットでは…	表示
#5036	個人認証を…	表示
…		

【図14】

[Figure 14]

図 1 4

